

Please notice that this is a DRAFT of a paper that has later been published as

Fossa, F. (2022). Autonomy and Automation. The Case of Connected and Automated Vehicles. In: P. Kommers, M. Macedo (eds.) *Proceedings of the International Conferences on ICT, Society, and Human Beings 2022; Web Based Communities and Social Media 2022; and E-Health 2022*, IADIS Press, pp. 244-248.

AUTONOMY AND AUTOMATION: THE CASE OF CONNECTED AND AUTOMATED VEHICLES

Fabio Fossa

*Department of Mechanical Engineering
Politecnico di Milano
Via Privata Giuseppe la Masa 1
20156 Milan (MI)
Italy*

ABSTRACT

This short paper offers a preliminary inquiry into the impacts of driving automation on personal autonomy. Personal autonomy is a key ethical value in western culture, and one that buttresses fundamental components of the moral life such as the exercise of responsible behaviour and the full enjoyment of human dignity. Driving automation simultaneously enhances and constrains it in significant ways. Hence, its moral profile with reference to the value of personal autonomy is uncertain. Ethical analysis shows that such uncertainty is due not just to the complexity of the technology, but also to the multifaceted normative profile of personal autonomy, which offers reasons to support both conditional and full driving automation. The paper sheds light on this duplicity, underlines the challenges this poses to the ethics of driving automation, and advocates for further research aimed at providing practitioners with more fine-grained guidelines on such a delicate issue.

KEYWORDS

Connected and Automated Vehicles; Automation; Personal Autonomy; Engineering Ethics; Ethics of Technology.

1. INTRODUCTION

The aim of this short paper is to shed light on how the ethical value of personal autonomy is impacted by driving automation. Its results, although preliminary, will hopefully contribute to raising awareness on the design and policy challenges that must be faced to effectively align future Connected and Automated Vehicles (CAVs) to such an important principle.

The quixotic relationship between personal autonomy and technological automation lies at the heart of several ethical quandaries across various AI-based applications (Laitinen & Sahlgren, 2021) such as, e.g., Autonomous Weapon Systems (Sharkey, 2019) and Recommender Systems (Varshney, 2020). Driving automation makes no exception (Chiodo, 2022). In this context, threats to and opportunities for personal autonomy are so numerous and deeply entangled with each other that much philosophical work is needed to clarify how personal autonomy is to be effectively pursued. The importance of such clarification should not be underestimated. Critically examining how widespread conceptions of personal autonomy apply to the case

of CAVs is key to realise relevant ethical opportunities and risks. Effective design choices and policy decisions importantly depend on it.

The paper is structured as follows. Section 2 discusses how the ethical value of personal autonomy has been brought to bear on driving automation. Section 3 analyses the definition of personal autonomy proposed in the European report *Ethics of Connected and Automated Vehicles* (Horizon, 2020) and draws attention to its composite nature. Applied to the field of driving automation, personal autonomy is accordingly specified as self-determination of driving tasks and freedom to pursue a good life through mobility. As Section 3 shows, however, these two specifications turn out to support different driving automation models – respectively: conditional and full automation – thus leaving doubts on how to practically comply with the demands of personal autonomy. Based on these results, Section 4 claims that the principle of personal autonomy requires further ethical elucidation in order to inspire unambiguous design and policy choices.

2. PERSONAL AUTONOMY AND DRIVING AUTOMATION

The importance of aligning the design, deployment, and use of CAVs to ethical expectations concerning personal autonomy could hardly be belittled. Even though its philosophical status is controversial (Magnani, 2020), personal autonomy enjoys widespread sociopolitical recognition as a key ethical value. In essence, it characterises human beings as self-determining entities who, therefore, deserve respect and protection (Christman, 2020). As such, it buttresses fundamental components of the moral life such as the exercise of responsible behaviour and the full enjoyment of human dignity. Given its relevance, personal autonomy evidently qualifies as a value to be pursued through technological innovation as well.

The European approach to the ethics of driving automation confirms the latter statement. In 2020, an interdisciplinary group of fourteen experts appointed by the European Commission authored the report *Ethics of Connected and Automated Vehicles. Recommendations on Road Safety, Privacy, Fairness, Explainability and Responsibility* (Horizon, 2020). The document establishes an ethical framework for CAVs and offers concrete recommendations aimed at guiding stakeholders in the effort of aligning driving automation to relevant ethical values. In close connection with the European approach to trustworthy Artificial Intelligence (AIHLEG 2019), the report starts by identifying and describing the basic normative cornerstones of the framework. Acknowledging its relevance, the authors indicate personal autonomy as one of the eight overarching ethical principles for driving automation along with non-maleficence, beneficence, dignity, responsibility, justice, solidarity, and inclusive deliberation (Santoni De Sio, 2021).

According to the report, the principle of personal autonomy states that human beings are to be conceived as “free moral agents” (Horizon, 2020, p. 22) whose right to self-determination ought to be respected. In relation to autonomous driving, personal autonomy demands that CAVs are designed so to “protect and promote human beings’ capacity to decide about their movements and, more generally, to set their own standards and ends for accommodating a variety of conceptions of a ‘good life’” (Horizon, 2020, p. 22). As such, autonomy plays a crucial role in several recommendations, ranging from the protection of privacy rights and the promotion of user choice to reducing opacity and enhancing explainability.

From an ethical point of view, the insistence on protecting and promoting personal autonomy in the context of driving automation seems appropriate. Bypassing individual decision-making through technical means risks leading to situations where personal decisions are taken by actors (e.g., designers, engineers, manufacturers, policy-makers) who, however, have no right nor particular competence to do so. This state of affair is evidently incompatible with the individual right to self-determination on personal matters and should be carefully avoided when designing or deploying CAVs.

The protection of autonomy in driving automation is also critical to support other important moral values. Considering human beings as free moral agents by principle means, at the same time, considering them responsible agents as well, to the extent that they can exercise such freedom. This is a necessary presupposition to establishing who is responsible, and why, when harmful consequences follow from the use of CAVs (Nyholm, 2018).

The value of personal autonomy, then, is vital to the ethics of driving automation for many reasons. On the one hand, CAVs designed and deployed in ways that promote personal autonomy will meet demands grounded on the protection of human dignity, thus supporting social acceptance and trust. On the other hand,

upholding personal autonomy is key to distributing responsibility in a clear and fair way while, at the same time, encouraging responsible behaviour. But how is personal autonomy to be pursued on a practical level?

3. ONE DEFINITION, TWO COMPONENTS

Whilst the ethical relevance of personal autonomy to driving automation is evident, it is difficult to specify how the principle is to be endorsed on a more tangible level. Driving automation, after all, consists in the delegation of driving tasks from human agents to digital systems. Arguably, constraints to personal autonomy are only to be expected. What needs to be further clarified, then, is how to automate driving functions without impacting too negatively on personal autonomy. This raises thorny practical questions. What aspects of human autonomy are relevant to driving automation? Which of them should be prioritised? What model of driving automation should be promoted through design and policy decisions? What guidelines should be offered to practitioners in this sense?

In order to answer these queries, a more fine-grained understanding is required of how personal autonomy is impacted by driving automation. In other words, it is necessary to identify which aspects of CAV users' experience qualify as expressions of their personal autonomy. These aspects, in turn, would serve as tangible constraints to driving automation: CAVs should be developed in ways that allow for their exercise. In sum, specifying how personal autonomy in driving automation actually looks like is a necessary step towards providing effective guidelines to stakeholders.

The definition of personal autonomy provided in the European report represents a good starting point to figure out more precisely what is at stake in this context. At a closer look, the definition exhibits two main components. They can be specified as (a) autonomy as self-determination of driving decisions; and (b) autonomy as freedom to pursue a good life through mobility.

(a) concerns the exercise of individual control over decisions that pertain to driving behaviour – i.e., to the ways in which the vehicle reaches its destination from its starting point. In the European report, this component is referred to when authors recommend to design, deploy, and use CAVs so to “protect and promote human beings' capacity to decide about their movements” (Horizon 2020, 22). In this sense, respecting CAV users' autonomy would mean to let them exercise some sort of control on the system operations that impact on their personal sphere.

(b), on the contrary, exhibits a wider scope. It refers to the freedom of pursuing happiness and to mobility as an important enabler of what makes life worth living. As stated in the European report, upholding personal autonomy also means to “protect and promote human beings' capacity to (...) set their own standards and ends for accommodating a variety of conceptions of a ‘good life’” (Horizon, 2020, p. 22). In this sense, aligning driving automation to the principle of personal autonomy would mean to envision CAVs as means to support the individual pursuit of personal flourishing and well-being.

In what follows, the ethical challenges related to complying with these two forms of personal autonomy are outlined. It is shown that complying with (a) would pose significant obstacles to compliance with (b), and vice versa. The paradoxical outcomes of the analysis suggest that further ethical research is needed to understand how personal autonomy can be pursued consistently in the context of driving automation.

4. A CONFLICT OF AUTONOMIES

Let us start by considering how personal autonomy as in (a) could be promoted through driving automation. In this sense, human autonomy partakes in driving automation mostly as a threatened individual value that requires to be adequately safeguarded. Particular care is required since the exercise of personal autonomy is variously constrained by driving automation (Xu, 2021). The experience of driving is a complex one, composed by a myriad of decisions, some of which are personal decisions or might have a considerable impact on the moral sphere. The delegation of such decisions to automated systems poses the risk of bypassing human judgment in ethically problematic ways.

Threats to the exercise of personal autonomy might variously arise in the context of driving automation. At low levels of automation, a speed control system that could not be overridden by human intervention even in case of emergency might be considered as problematic with reference to personal autonomy

(Schoonmaker, 2016). At the opposite extreme, suppose that full autonomous vehicles will be able to distribute harm during unavoidable collisions according to given ethical values. In this case, it might be problematic in terms of personal autonomy if said values were set not by passengers themselves, but rather by other stakeholders (Millar, 2016; Contissa *et al.*, 2017; Millar, 2017). Considering less futuristic scenarios involving high levels of automation, automated features concerning ethical driving behaviour – e.g., regarding the safety distance to be accorded to vulnerable road users or traffic etiquette at pedestrian crossings – might qualify as constraints to the exercise of personal autonomy. Finally, relying on CAVs would restrict the possibility of taking timely decisions concerning routes, which could variously impact the execution of self-determined intentions – e.g., staying away from given roads to protect one’s privacy (Boeglin, 2015).

In light of the above, it seems reasonable to conclude that the rush towards full automation should not hinder limited forms of human control over driving tasks, at least when this would serve the legitimate expression of personal autonomy. As suggested, for instance, by the Meaningful Human Control approach (Santoni de Sio & van den Hoven, 2018), if some driving decisions are for users to make, then CAVs should allow for their personal autonomy to be expressed. Arguably, in a context of full automation this could only be accomplished indirectly, e.g., through the setting of user preferences. It is at least uncertain, however, whether this form of indirect control over system operations would satisfy the demands of the principle of personal autonomy. More likely, (a) seems to encourage the development of automated features that leave enough space for the exercise of user autonomy – as happens in conditional automation, where control over driving tasks is shared with the system rather than fully delegated to it.

The claim according to which personal autonomy would be better served by conditional automation is controversial, however, if the value is intended as in (b). Driving automation can indeed have beneficial impacts on human autonomy as the freedom to pursue a good life. At least two opportunities stand out: inclusive transportation and the reappropriation of travel time. Both importantly enable the possibility to fulfil personal needs and desires, thus increasing well-being.

On the one hand, CAVs could massively improve the autonomy of social categories that are currently excluded from manual driving because of physical and cognitive impairments. Independent access to transportation is critical for pursuing personal well-being and leading a satisfying social life. Since driving tasks would be automated, physical and cognitive impairments would no longer constitute an insurmountable barrier to the autonomous use of road vehicles (Lim & Tacihagh, 2018). On the other hand, driving automation could support the self-determined pursuit of a good life by allowing users to reclaim travel time. Freed from the burden of driving themselves, CAV users would be able to employ travel time as they prefer. In addition, autonomous decision-making on matters that importantly impact on individual well-being would also be supported. For instance, decisions about where to live would be less constrained by work locations and other circumstantial factors.

In both of the above cases, personal autonomy benefits entirely depend on full automation. As a matter of fact, individuals excluded from manual driving would be poor candidates for shared control as well (Goggin, 2019). Similarly, full delegation is necessary for CAV users to freely engage in other, more satisfying activities. In order to support autonomy as freedom to pursue one’s own conception of a good life, then, human intervention and supervision should be increasingly automated away.

The contrast between a) and b) is evident. Supporting both partial and full automation, compliance with the ethical principle of personal autonomy steers in directions that are difficult to harmonise. Analogously, it is hard to realise how protecting the exercise of user self-determination over driving decisions can go hand in hand with protecting the right to a self-determined good life pursued through mobility. This ambiguity, that stems from the complexity of the notion of autonomy and competing expectations about driving automation, represents a barrier towards designing CAVs that protect and promote personal autonomy. Uncertainty on this matter leaves engineers with the puzzling task of figuring out in what sense personal autonomy can be a value to embed in driving automation, or how to do so.

5. CONCLUSION

When human autonomy meets driving automation, two of its essential components come into conflict. Fully delegating driving to CAVs would limit (or entirely bypass) personal autonomy as the self-

determination of driving decisions, so that conditional automation appears to be the most promising option. However, autonomy as the pursuit of self-determined life preferences, interests, goals, and values is best supported by full automation.

The tension that obtains poses a most delicate issue to the ethics of driving automation. Aligning future CAVs to ethical expectations in terms of personal autonomy is an important task. Infringements in this sense are likely to generate distrust and public backlash. However, the paradoxical nature of the issue makes it complicated to move from abstract endorsements to more practical design and policy recommendations.

Future philosophical research must tackle this obscurity and provide less ambiguous accounts of personal autonomy in the context of driving automation. Meanwhile, ambiguity must be assumed as a given. Learning how to deal with it is then of utmost importance. The most urgent task on a design and policy level, then, likely consists in promoting reflection on possible threats to personal autonomy – however ill-defined the concept might be – and assessing solutions aimed at minimising potential harm. Such preliminary, applied ethics work might in turn offer precious help to further refine the notion of personal autonomy as it applies to driving automation (Fossa *et al.*, 2022).

To conclude, shedding light on what it means for driving automation to comply with the value of personal autonomy reveals a series of complicated issues that calls for further analysis. As a first step in this direction, the present paper has offered a preliminary contribution to the identification of such challenges and their origin. By doing this, it has set the stage for future research aimed at better defining the conceptual profile of personal autonomy as it concerns the ethics of driving automation.

REFERENCES

- Boeglin, J.A., 2015. The Costs of Self-Driving Cars: Reconciling Freedom and Privacy with Tort Liability in Autonomous Vehicle Regulation. *Yale Journal of Law and Technology*, Vol. 17, No. 4, pp. 171-203.
- Chiodo, S., 2022. Human autonomy, technological automation (and reverse). *AI & Society*, Vol. 37, pp. 39-48. <https://doi.org/10.1007/s00146-021-01149-5>
- Christman, J., 2020. Autonomy in Moral and Political Philosophy. In: Zalta, E. (ed.), *The Stanford Encyclopedia of Philosophy (Fall 2020 Edition)*. <https://plato.stanford.edu/archives/fall2020/entries/autonomy-moral/>
- Contissa, G., *et al.*, 2017. The Ethical Knob: ethically-customisable automated vehicles and the law. *Artificial Intelligence and Law*, Vol. 25, pp. 365-378. <https://doi.org/10.1007/s10506-017-9211-z>
- Fossa, F., *et al.*, 2022. Operationalizing the Ethics of Connected and Automated Vehicles: an Engineering Perspective. *International Journal of Technoethics*, Vol. 13, No. 1, pp. 1-20. <https://doi.org/10.4018/IJT.291553>
- Goggin, G., 2019. Disability, Connected Cars, and Communication. *International Journal of Communication*, Vol. 13, pp. 2748-2773. <https://ijoc.org/index.php/ijoc/article/view/9021>
- HLEGAI – High-Level Expert Group on Artificial Intelligence, 2019. *Ethics Guidelines for Trustworthy AI*. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (last accessed May 3rd, 2022).
- Horizon 2020 Commission Expert Group to advise on specific ethical issues raised by driverless mobility (E03659), 2020. *Ethics of Connected and Automated Vehicles: recommendations on road safety, privacy, fairness, explainability and responsibility*. <https://op.europa.eu/en/publication-detail/-/publication/89624e2c-f98c-11ea-b44f-01aa75ed71a1/language-en> (last accessed May 3rd, 2022).
- Laitinen, A., and Sahlgren, O., 2021. AI Systems and Respect for Human Autonomy. *Frontiers in Artificial Intelligence*, Vol. 4, 705164, pp. 1-14. <https://doi.org/10.3389/frai.2021.705164>
- Lim, H.S.M., and Taeihagh, A., 2018. Autonomous Vehicles for Smart and Sustainable Cities: An In-Depth Exploration of Privacy and Cybersecurity Implications. *Energies*, Vol. 11, No. 5, 1062. <https://doi.org/10.3390/en11051062>
- Magnani, L., 2020. Autonomy and the Ownership of Our Own Destiny: Tracking the External World and Human Behavior, and the Paradox of Autonomy. *Philosophies*, Vol. 5, No. 53, pp. 1-12. <https://doi.org/10.3390/philosophies5030012>
- Millar, J., 2016. An Ethics Evaluation Tool for Automating Ethical Decision-Making in Robots and Self-Driving Cars. *Applied Artificial Intelligence*, Vol. 30, No. 8, pp. 787-809. <http://dx.doi.org/10.1080/08839514.2016.1229919>
- Millar, J., 2017. Ethics Settings for Autonomous Vehicles. In: Lin, P., Abney, K., Jenkins, R. (eds.), *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. Oxford University Press, Oxford, pp. 20-34. <https://doi.org/10.1093/oso/9780190652951.003.0002>
- Nyholm, S., 2018. The ethics of crashes with self-driving cars: a roadmap II. *Philosophy Compass*, Vol. 13, No. 7, e12506. <https://doi.org/10.1111/phc3.12506>

- Santoni de Sio, F., 2021. The European Commission report on ethics of connected and automated vehicles and the future of ethics of transportation. *Ethics and Information Technology*, Vol. 23, pp. 713-726. <https://doi.org/10.1007/s10676-021-09609-8>
- Santoni De Sio, F., and van den Hoven, J., 2018. Meaningful human control over autonomous systems: A philosophical account. *Frontiers of Robotics and AI*, Vol. 5, No. 15, pp. 1-14. <https://doi.org/10.3389/frobt.2018.00015>
- Schoonmaker, J., 2016. Proactive privacy for a driverless age. *Information & Communications Technology Law*, pp. 1-33. <https://doi.org/10.1080/13600834.2016.1184456>
- Sharkey, A., 2019. Autonomous weapons systems, killer robots and human dignity. *Ethics and Information Technology*, Vol. 21, pp. 75-87. <https://doi.org/10.1007/s10676-018-9494-0>
- Varshney, L.R., 2020. Respect for Human Autonomy in Recommender Systems. *ArXiv*, *abs/2009.02603*. <https://arxiv.org/abs/2009.02603v1>
- Xu, W., 2021. From Automation to Autonomy and Autonomous Vehicles. *Interactions*, January-February, pp. 49-53. http://dl.acm.org/ft_gateway.cfm?id=3434580&type=pdf&dwn=1